US (HRS):

We are using the Harmonized updated version C, until 2019 that was updated now in 2022 and contains 42,233 observations. It is a Respondent level file so each row represents a unique Respondent. This leaves us with 18,747 observations using only wave 12 (year 2014) of HRS. It is still important to use the updated version as they constantly correct the harmonized variables.

Mexico MHAS:

Version B.4 incorporates the latest released version of MHAS data, and adds several new variables. It contains 22,016 observations or rows- 22016. We are using the Harmonized VERSION B.4 (2001-2015), February 2022, for the MHAS data. The Mexican Health and Aging Study (MHAS) is a longitudinal household survey dataset for the study of health, economic position, and quality of life among the elderly. MHAS datasets as of September 2020. The MHAS (Mexican Health and Aging Study) Version B.4 incorporates the latest released version of MHAS data, and adds several new variables. It contains 22,016 observations or rows. It is a Respondent-level file so each row represents a unique Respondent.

We will focus on Wave 4, which is for years 2014/2015. We will have 17,616 observations.

England (ELSA):

We are using the Version G.2 (2002-2019), July 2021 for The English Longitudinal Study on Ageing (ELSA). It is a longitudinal household survey dataset for the study of health, economic position, and quality of life among the elderly (panel survey of people aged 50 and over and their partners, living in private households in England). Version G.2 incorporates the latest released version of ELSA data, which includes eleven main modules and the associated datasets, and adds variables and observations from Wave 9 with a total of 19,802 observations. It also adds new variables and makes adjustments and corrections. We will focus on Wave 7, nonetheless. The samples have been drawn from households which previously responded to the Health Survey for England (HSE). The seventh wave was conducted between June 2014 and May 2015 and included a refreshment sample selected from HSE 2011-2012.

Acknowledgements. “This analysis uses data or information from the Harmonized ELSA dataset and Codebook, Version G.2 as of July 2021 developed by the Gateway to Global Aging Data. The development of the Harmonized ELSA was funded by the National Institute on Aging (R01 AG030153, RC2 AG036619, R03 AG043052). For more information, please refer to <https://g2aging.org/>.”

India (LASI)

The Longitudinal Aging Study in India (LASI) is a multidisciplinary, internationally harmonized panel study designed to be nationally representative of India’s population aged 45 and older. LASI is a joint project of three partnering institutions: International Institute for Population Sciences (IIPS), Harvard T.H. Chan School of Public Health (HSPH), and University of Southern California (USC). The first wave was conducted between 2017 and 2019 in 35 of India’s 36 states and union territories (except Sikkim). This initial sample, as released by USC, included 42,951 households and 72,262 individuals. The LASI sampling plan is complex and was based on the 2011 Indian Census with a multistage, stratified cluster sample design. The sample design includes three distinct selection stages in rural areas and four stages in urban areas. We use Version A.2 that makes corrections using the January 2021 released version of Wave 1 of the LASI data.

Europe- SHARE

This is Version F in the harmonized files and incorporates the latest released version of SHARE data, release 8.0.0, and adds observations from Wave 8. It contains 139,620 observations or rows. It is a Respondent-level file so each row represents a unique Respondent. It also adds new variables and makes adjustments and corrections. We focus on data from SHARE Wave 6, with the release 8.0.0 as of February 2022. SHARE uses a multistage stratified sample. Its weighting variables make its data representative of the target populations in constituent countries.

Wave 6 does not still have full coverage of European countries, with the following countries only added in Wave 7: Finland, Lithuania, Latvia, Slovakia, Romania, Bulgaria, Malta and Cyprus. The final sample size is 66,708 observations for the pooled countries after the missings are excluded.

Acknowledgements. "This analysis uses data or information from the Harmonized SHARE dataset and Codebook, Version F as of June 2022 developed by the Gateway to Global Aging Data. The development of the Harmonized SHARE was funded by the National Institute on Aging (R01 AG030153, RC2 AG036619, R03 AG043052). For more information, please refer to “https://g2aging.org/”.

China -CHARLS

The China Health and Retirement Longitudinal Study (CHARLS) is a longitudinal study of individuals over age 45 in China. Version D incorporates the latest released version of CHARLS data, and adds variables for Wave 4. It contains 25,586 observations or rows. It is a Respondent-level file so each row represents a unique Respondent; The sample population was selected as part of a stratified, multistage probability design. We will use Wave 3. As we concentrate on ages 50 and above due to the other samples we do not include individuals younger than 50. This leaves us with a sample size of 16,344 individuals.

Acknowledgements . "This analysis uses data or information from the Harmonized CHARLS dataset and Codebook, Version D as of June 2021 developed by the Gateway to Global Aging Data. The development of the Harmonized CHARLS was funded by the National Institute on Aging (R01 AG030153, RC2 AG036619, R03 AG043052). For more information, please refer to <https://g2aging.org/>.” This document used CHARLS Waves 1 through 4 as of June 2021. CHARLS is supported by Peking University, the National Natural Science Foundation of China, the National Institute on Aging, and the World Bank.

KLOSA- Korea

The Korean Longitudinal Study of Ageing (KLoSA) is a panel survey of people aged 45 and over and their partners, living in private households in Korea. The survey elicits information about demographics, income, assets, health, cognition, family structure and connections, health care use and costs, housing, job status and history, expectations, and insurance.

KLoSA surveys respondents every two years. Funded by the Korean Ministry of Labor, the Korean Institute of Labor (KLI) collected the first two waves, and the Korea Employment Information Service (KEIS) collected the Waves 3, 4, 5 and 6 of KLoSA, with the first wave of the KLoSA survey being conducted in fall/winter of 2006. The sample population was selected as part of a stratified, multi-stage area probability design. The first component of this sampling framework is the probability proportional to size (PPS) systematic sampling of the 2005 (South Korean) Census enumeration districts after stratifying by the location (15 major metropolitan cities and provinces) and characteristic of the district (urban or rural, and apartment building or non-apartment dwelling). Households were selected within PSUs from a listing of households in the Census identified as age-eligible; that is, inhabited by at least one person 45 years of age and older. This initial sample included 10,254 respondents age 45 and over. The second wave was conducted in 2008 and had 8,688 respondents. The third wave was conducted in 2010 and had 7,920 respondents. The fourth wave was conducted in 2012 and had 7,486 respondents. There was no refresher sample in Waves two through four. In 2014, a refreshment sample of individuals born in 1962 or 1963 was drawn and it included 920 individuals, which were added to the 7,029 remaining core sample respondents for a total of 7,949 Wave 5 respondents. The sixth wave was conducted in 2016 and had 7,490 respondents. We will focus on Wave 5. However, because we focus on ages >50, the sample is not 7,949, but . We use the hamornized Version C that contains 11,174 observations or rows. It is a Respondent-level file so each row represents a unique Respondent. It also adds new variables and makes adjustments and corrections.

Education as the most important driver of sex differences in health. Is there a relationship between compression, expansion and dynamic equilibrium and inequality?